BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

from prokaryotic or eukaryotic sources. Usually, the polypeptide will have at least 15 amino acids (gene of 45 bp), more usually 30 amino acids (gene of 90 bp), and may be 300 amino acids (gene of 900 bp) or greater.

Polypeptides of interest include enzymes, surface membrane proteins, mammalian proteins, such as those involved in regulatory functions, such as lymphokines, growth factors, hormones, etc., blood clotting factors, clot degrading factors, immunoglobulins, etc.

One of the methods for selectable cleavage is cyanogen bromide which is described in U.S. Patent No. 4,366,246. This technique requires the absence of an available methionine other than at the site of cleavage or the ability to selectively distinguish between the methionine to be cleaved and a methionine within the polypeptide sequence. Alternatively, one may use various proteases which recognize specific amino acid sequences, such as the serine proteases, illustrated by trypsin, pepsin, bromelain, papain, or the like, where a series of basic amino acids, such as lysine and arginine act as a cleavage site. Peptidases can be employed which are specific for particular sequences, such as those peptidases which are involved in the selective cleavage of secretory leader signals from a polypeptide. These enzymes are specific for such sequences which are found with α-factor and killer toxin in yeast, such as KEX 2 endopeptidase with specificity for pairs of basic residues (Julius et al., Cell (1984) 37:1075-1089). Other enzymes which recognize and cleave specific sequences include: bovine enterokinase (Light et al., Anal. Biochem. (1980) 106:199-206); collagenase (Germino and Batia, Proc. Natl. Acad. Sci. (1984) 81:4692-4696); factor X (Nagai & Thøgersen, Nature (1984) 309:810-812); polyubiquitin processing enzyme (Ozakaynak et al., Nature (1984) 312:663-666).

Where one or more amino acids are involved in the cleavage site, the codons coding for such sequence

EXHIBIT B

may be prepared synthetically and ligated to the sequences coding for the polypeptides so as to provide for a fused protein where all the codons are in the proper reading frame and the selectable cleavage site joins the two polypeptides.

Instead of only a small portion of the fused coding sequence being synthetically prepared, the entire sequence may be synthetically prepared. This allows for certain flexibilities in the choice of codons, whereby one can provide for preferred codons, restriction sites, avoid or provide for particular internal structures of the DNA and messenger RNA, and the like.

While for the most part, the fused coding sequence will be prepared as a single entity, it should be appreciated that it may be prepared as various fragments, these fragments joined to various untranslated regions, providing for particular functions and ultimately the coding sequences brought together at a subsequent stage. However, for clarity of presentation, the discussion will be directed primarily to the situation where the coding sequence is prepared as a single entity and then transferred to an expression vector.

The various sequences comprising the parts of the fused coding sequence can be joined by introducing a first fragment into a cloning vector. The resulting clone may then be restricted at a site internal to the coding sequence and an adapter introduced which will replace any lost codons and which has a convenient terminus for joining to the next fragment. The terminus may be cohesive or blunt-ended, depending upon the particular nucleotides involved. After cloning of the combined first fragment and adapter, the vector may be restricted at the restriction site provided by the adapter and the remaining coding sequence of the second fragment introduced into the vector for ligation and cloning. The resulting fused sequence should be flanked

polypeptide. Various secretory leader and processing signals are known, such as yeast α-factor, yeast killer toxin and the like. The DNA sequence coding for these polypeptide signals may be linked in proper reading frame to the 5'- end (in direction of transcription of the sense strand) of the DNA sequence coding for the fused polypeptide to provide for transcription and translation of a pre-fused polypeptide.

In accordance with the subject invention, the product is produced in at least a 5 weight percent, preferably at least 6 weight percent, and more preferably at least about 10 weight percent, of the total protein of the host. In this manner, the nutrients employed are efficiently utilized for conversion to a desired product.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

CONSTRUCTION OF EXPRESSION VECTORS

Construction of pYSI1

A yeast expression plasmid pYSI1, containing the human SOD gene fused to the amino-terminus of human proinsulin gene, under the regulation of the GAP promoter and terminator was constructed. A triplet coding for methionine was included between the SOD and proinsulin genes to allow for chemical processing of the fusion protein. The SOD sequences correspond to a cDNA isolated from a human liver library, except for the first 20 codons which were chemically synthesized. The proinsulin sequence was chemically synthesized according to the amino acid sequence reported by (Bell et al. (1979), Nature 282:525-527), but using yeast preferred codons. The GAP promoter and terminator sequences were obtained from the yeast GAP gene (Holland & Holland, J. Biol. Chem. (1979) 254:5466-5474) isolated from a yeast li-

Plasmid pYSI1 was constructed as follows. Three fragments were employed which involve a 454bp NcoI-Sau3A isolated from phSOD (also designated as pSODNco5), where the fragment includes the entire coding sequence for human superoxide dismutase (hSOD) with the exception of the last three 3'- codons; a 51bp Sau3A-HindIII synthetic adapter, which codes for the last three codons of hSOD, methionine, and the first 14 codons of proinsulin; and a 231bp HindIII-SalI fragment, isolated from pINS5, which encodes proinsulin excepting the first 14 amino acids. These fragments were ligated together and introduced into the plasmid pPGAP, which had been previously digested with NcoI and SalI and alkaline phosphatase treated. The resulting plasmid pSI1 was digested with BamHI to provide an expression cassette which was cloned into plasmid pC1/1 to yield pYSI1.

Plasmid phSOD (also designated as pSODNco5) is a pBR322-derived bacterial expression vector which contains a complete cDNA coding (except that the first 20 codons were chemically synthesized) for hSOD as described in copending application serial number 609,412 filed on May 11, 1984. Plasmid pINS5 is a pBR322-derived vector which contains a proinsulin coding sequence chemically synthesized according to the amino acid sequence reported by Bell et al., Nature (1979) 282:525-527. Plasmid pPGAP is a pBR322-derived vector described in copending application 609,412 (supra) which contains a GAP promoter and GAP terminator (Holland and Holland, J. Biol. Chem. (1979) 254:5466-5474) with a polylinker between them, which provides for single restriction sites for cloning. Plasmid pC1/1 is a yeast expression vector which includes pBR322 sequences, 2μ plasmid sequences and the yeast gene LEU2 as a selectable marker. See EPO 83/306507.1, which relevant parts are incorporated herein by reference.

Construction of pYS12

To prepare the fused gene having the hSOD coding sequence at the 3'-terminus in the direction of transcription separated from the proinsulin gene by a "spacer" of codons coding for K-R-S-T-S-T-S, the following fragments were ligated. A 671bp BamHI-SalI fragment containing the GAP promoter, the proinsulin gene and codons for the spacer amino acids; a 14bp SalI-NcoI synthetic adapter, which codes for the last spacer amino acids as a junction of both genes; and a 1.5kb NcoI-BamHI fragment isolated from pC1/1 GAPSOD described in copending application 609,412 (supra), which includes the hSOD coding region, 56bp of hSOD terminator and 934bp of GAP terminator region. The resulting cloned fragment was isolated and inserted into BamHI digested, alkaline phosphatase treated pC1/1.

Plasmids pPKI1 and pPKI2.

Plasmids homologous to pYSI1 and pYSI2, but using the yeast pyruvate kinase (PYK) gene instead of hSOD gene, were also constructed. pPKI1 contains the PYK coding sequence fused to the amino-terminus of the human proinsulin gene under regulation of the yeast PYK promoter and yeast GAP terminator. pPKI2 contains the PYK coding sequence at the 3'-terminus in the direction of transcription separated from the proinsulin gene by a "spacer" of codons coding for K-R-S-T-S. This fused gene is under regulation of the GAP promoter and PYK terminator.

Construction of pYASI1

This yeast expression plasmid is similar to pYSI1 and contains the hSOD gene fused to the amino terminus of the human proinsulin gene, with a methionine codon at the junction between both genes. The fusion

was constructed by replacing the GAP promoter sequence from pYSI1 with the hybrid ADH2-GAP promoter sequence.

The ADH2 portion of the promoter was constructed by cutting a plasmid containing the wild type ADH2 gene (plasmid pADR2, see Beier and Young, Nature (1982) 300:724-728) with the restriction enzyme EcoR5, which cuts at a position +66 relative to the ATG start codon, as well as in two other sites in pADR2, outside of the ADH2 region. The resulting mixture of a vector fragment and two smaller fragments was resected with Bal31 exonuclease to remove about 300bp. Synthetic XhoI linkers were ligated onto the Bal31 treated DNA. The resulting DNA linker terminator fragment was separated from the linkers by column chromatography, cut with the restriction enzyme XhoI, religated and used to transform E. coli to ampicillin resistance. The positions of the XhoI linker additions were determined by DNA sequencing. One plasmid which contained an XhoI linker located within the 5' non-transcribed region of the ADH2 gene (position -232 from ATG) was cut with the restriction enzyme XhoI, treated with nuclease S1, and subsequently treated with the restriction enzyme EcoRI to create a linear vector molecule having one blunt end at the site of the XhoI linker and an EcoRI end.

The GAP portion of the promoter was constructed by cutting plasmid pPGAP (supra) with the enzymes BamHI and EcoRI, followed by the isolation of the 0.4Kbp DNA fragment. The purified fragment was cut with the enzyme AluI to create a blunt end near the BamHI site.

Plasmid pJS014 was constructed by the ligation of the AluI-EcoRI GAP promoter fragment to the ADH2 fragment present on the linear vector described above.

Plasmid pJS104 was digested with BamHI (which cuts upstream of the ADH2 region) and with NcoI (which cuts downstream of the GAP region). The about 1.3Kbp

purified and ligated to an about 1.7Kbp fragment containing the hSOD-proinsulin fusion DNA sequences and GAP terminator present in pYSI1 (previously described). This 3Kbp expression cassette was cloned into BamHI digested and phosphatase treated pC1/1 to yield pYASI1.

Expression of fusion proteins

Yeast strain 2150-2-3 (Mata, ade 1, leu 2-04, cir°) was transformed with the different vectors according to Hinnen et al., Proc. Natl. Acad. Sci. USA (1978) 75:1929-1933. Single transformant colonies harboring constitutive GAP regulated vectors were grown in 2ml of leu$^-$ selective media to late log or stationary phase. Cells harboring inducible ADH2-GAP and regulated vectors were grown to saturation in leu$^-$ selective media, subsequently diluted 1:20 (v/v) in YEP, 3% ethanol, and grown to saturation in this medium. Cells were lysed in the presence of SDS and reducing agent and the lysates clarified by centrifugation. Cleared lysates were subjected to polyacrylamide gel electrophoresis (Laemmli, Nature (1970) 277:680). Following staining with Coomassie blue, a band of about 28kDal (kilodaltons) was observed, the size predicted for the fusion protein. This band was detected in those cells transformed with expression vectors, while being absent from extracts of cells harboring control (pC1/1) plasmids. The fusion protein accounts for over 10% of the total cell protein as estimated from the stained gels in those cells transformed with pYSI1, pYSI2 or pYASI1, while it accounts for less than 0.5% in pYPKI1 or pYPKI2 transformants. The following Table indicates the results:

Plasmid	Description of sequences contained in the expression cassette.	Expression (percent of total cell protein)
pYPKI 1	PYK_p PYK M BCA5 GAP_t	0.5%
pYPKI 2	GAP_p M BCA5 ~~KR(ST)$_2$~~ KRSTS PYK PYK_t	0.5%
pYSI 1	GAP_p SOD M BCA5 GAP_t	10 %
pYSI 2	GAP_p M BCA5 ~~KRSTS~~ KR(ST)$_2$S SOD GAP_t	10 %
pYASI 1	$(ADH2\text{-}GAP)_p$ SOD M BCA5 GAP_t	10 %

Note: Proinsulin accounts for less than 0.1% of total cell protein in cells transformed with pYGAPINS5, a plasmid containing the proinsulin gene under regulation of GAPDH promoter and terminator (GAP_p M BCA5 GAP_t).

* PYK : pyruvate kinase gene
SOD : human SOD gene
BCA5: proinsulin gene
M, K, R, S, T: one letter amino acid code

GAP_p : GAP promoter
GAP^t : GAP terminator
PYK^t : PYK promoter
PYK^p : PYK terminator
$(ADH2\text{-}GAP)_p$: hybrid ADH2-GAP promoter

Results shown in the Table indicate that while expression levels of PYK-proinsulin fusion are comparable to those obtained with proinsulin alone (about 0.5% and 0.1%, respectively), the expression levels of hSOD-proinsulin are about 20 to 100 fold higher. The inducible ADH2-GAP hybrid transcriptional initiation regulatory region is preferred, since it is noted that constitutive production in scaled-up cultures results in unstable expression.

The hSOD-proinsulin proteins synthesized by yeast were also submitted to Western analysis. Cleared yeast lysates prepared as described above were electrophoresed on polyacrylamide gels (Laemmli, supra) and proteins were subsequently electroblotted onto nitrocellulose filters (Towbin et al., Proc. Natl. Acad. Sci. USA (1979) 76:3450). Two identical filters were blotted.

guinea pig anti-insulin antibodies for 12hr at 4°C. Both sera had been preadsorbed with pC1/1 control lysate in 10% goat serum. The filters were washed with 1% BSA PBS and a second goat anti-rabbit or anti-guinea pig antibody conjugated with horseradish peroxidase added. Finally, the filters were incubated with horseradish peroxidase color development reagent (Bio-Rad) and washed. The Western analysis showed that the fusion protein reacted with both antibodies.

Cleavage of the fusion proteins

A saturated culture of 2150 (pYASI1) was grown in SDC minus leucine plus threonine and adenine, containing 2% glucose. This was used to inoculate a 10 liter fermentor containing YEP with 3% ethanol as carbon source. After 48hrs at 30°C, the cells were harvested by centrifugation (Sharples), weighed (124g), and washed with cold water.

The cells were lysed by glass bead disruption (Dyno mill) using a buffer containing 10mM Tris Cl, pH 7.0, 1mM EDTA, 1µg/ml pepstatin A and 1mM PMSF. The mixture was centrifuged for 20min at 8,000rpm in a JA10 rotor (Beckman). The pellet was resuspended in 100mls of buffer and the liquid was removed from the beads. This was repeated until ~500mls of buffer was used to thoroughly remove all pellet material from the glass beads. The resuspended pellet was centrifuged, and the pellet washed a second time. The pellet was then extracted for 30min in buffer plus 1% SDS.

The SDS soluble fraction was ion-pair extracted using 500mls of solvent A (Konigsberg and Henderson, (1983) Meth. in Enz. 91, pp. 254-259), the pellet washed once with solvent A, and once with acetone.

After drying in a vacuum dessicator, the powder was dissolved in 140mls 100% formic acid. Sixty

was added, and the reaction continued for 24hrs. At this time, the material was dialyzed overnight against 4 liters H_2O using 2000 MW cutoff tubing (Spectrapor). A second dialysis against 0.1% acetic acid followed. After lyophilization, a powder consisting mostly of SOD-homoserine lactone and proinsulin was obtained, weighing 1.1g.

This powder was dissolved in a 200ml solution of 7% urea, 9% sodium sulfite, and 8.1% sodium tetrathionate - $2H_2O$, pH 7.5. After incubation for 3hrs at 37°C, the S-sulfonate products were dialyzed twice versus 10mM Tris pH 8.0, and once versus 20mM TEAB (triethylammonium bicarbonate), pH 7.3.

The S-sulfonates were recovered by lyophilization and dissolved in 240mls DEAE column buffer (Wetzel et al., Gene (1981) 17:63-71) and loaded onto a 60ml column. After washing with two column volumes, the proinsulin-S-sulfonate was eluted with a 600ml gradient of 0 to 0.4M NaCl in column buffer. Fractions containing proinsulin S-sulfonate were pooled and dialyzed twice against 10mM Tris, pH 7.5, and once against 1mM Tris.

The product, ~90% pure proinsulin-S-sulfonate, was shown to migrate at expected on pH 9 gel electrophoresis (Linde et al., Anal. Biochem. (1980) 107:165-176), and has the correct 15 N-terminal residues. On analysis, the amino acid composition was very close to that expected, not exactly correct due to the presence of a low level of impurities. The yield was 150mg.

Preliminary results have been obtained with the following procedural. The proinsulin-S-sulfonate can be renatured at pH 10.5, with β-mercaptoethanol (Frank et al. (1981) in Peptides: Synthesis, Structure and Function, Proceedings of the Seventh American Peptide Symposium, Rich and Gross, eds., Pierce Chemical Co., Rockford, IL, pp. 729-738). In preliminary experi-

WHAT IS CLAIMED IS:

1. In a method for preparing a polypeptide in a eukaryotic host, where the polypeptide is heterologous to the host and is expressed in low percentage amounts of total protein, the improvement which comprises:

joining an open reading frame DNA sequence coding for said polypeptide with a second open reading frame DNA sequence coding for a heterologous polypeptide produced in high yield and stable in said eukaryotic host, where the two polypeptides are joined to form a selectively cleavable link or the two open reading frames are joined by a sequence coding for a selectively cleavable link, to form a fusion polypeptide;

introducing the sequence coding for said fusion polypeptide under conditions for expression in said eukaryotic host, whereby said fusion polypeptide is expressed; and

isolating and cleaving said fusion polypeptide to provide said second polypeptide in high yield.

2. A method according to claim 1, wherein said cleavable link is enzymatically cleavable.

3. A method according to claim 1, wherein said cleavable link is a methionine.

4. A method according to claim 1, wherein said eukaryotic host is yeast.

5. A method according to claim 4, wherein said DNA sequences are under the transcriptional regulatory control of a transcriptional initiation regulatory region comprising a promoter region for a glycolytic

6. A method according to claim 5, wherein said transcriptional initiation regulatory region is inducible.

7. In a method for preparing a mammalian polypeptide in a yeast host, where the polypeptide is expressed in low percentage amounts of total protein, the improvement which comprises:

joining an open reading frame DNA sequence coding for said polypeptide with a second open reading frame DNA sequence coding for heterologous superoxide dismutase (SOD), where the two polypeptides are joined to form a selectively cleavable link or the two open reading frames are joined by a sequence coding for a selectively cleavable link, to form a fusion polypeptide;

introducing the sequence coding for said fusion polypeptide under conditions for expression in said yeast, whereby said fusion polypeptide is expressed; and

isolating and cleaving said fusion polypeptide to provide said mammalian polypeptide in high yield.

8. A method according to claim 7, wherein said superoxide dismutase is human superoxide dismutase (hSOD).

9. A method according to claim 7, wherein said mammalian polypeptide is 3' to said SOD coding sequence in the direction of transcription.

10. A method according to claim 7, wherein said mammalian polypeptide is 5' to said DNA sequence coding for SOD in the direction of transcription.

11. A method according to claim 7, wherein

12. A method according to claim 11, where said transcriptional initiation regulatory region consists essentially of a glycolytic enzyme promoter region and ADH2 control region.

13. A DNA sequence coding for superoxide dismutase joined to a DNA sequence coding for a mammalian polypeptide, where the two coding sequences are joined by bases coding for amino acids defining a selectively cleavable site or a link having at least one amino acid, wherein said link provides for a selectively cleavable site.

14. A DNA sequence according to claim 13, wherein said link codes for methionine.

15. A DNA sequence according to claim 13, wherein said link codes for an enzymatically removable link.

16. An expression sequence including in direction of transcription, an inducible transcriptional initiation regulatory region and a DNA sequence according to claim 13.

17. A polypeptide encoded for by a DNA sequence according to claim 13.

18. A polypeptide according to claim 17, wherein said mammalian polypeptide encodes for at least a portion of proinsulin.